

## **Report for Grant NAG5-13385**

Submitted to:  
**National Aeronautic and Space Administration**  
Grants Office  
Goddard Space Flight Center  
Greenbelt, MD 20771

July 2005

NASA Technical Contact:  
**James Fischer**  
GSFC Code 930

Submitted by:  
**Phillip Merkey**  
Department of Computer Science  
Michigan Technological University

### **Abstract**

This grant supported the effort to characterize the problem domain of the Earth Science Technology Office's Computational Technologies Project, to engage the Beowulf Cluster Computing Community as well as the High Performance Computing Research Community so that we can predict the applicability of said technologies to the scientific community represented by the CT project and formulate long term strategies to provide the computational resources necessary to attain the anticipated scientific objectives of the CT project. Specifically, the goal of the evaluation effort is to use the information gathered over the course of the Round-3 investigations to quantify the trends in scientific expectations, the algorithmic requirements and capabilities of high-performance computers to satisfy this anticipated need.

## Introduction

The overall goal of ESTO's Computational Technologies Project is to demonstrate the potential afforded by a balanced teraflop system's performance to further our understanding of and ability to predict the dynamic interaction of physical, chemical and biological processes affecting the Earth, the solar-terrestrial environment, and the universe. In order to successfully demonstrate this potential we must not only develop Grand Challenge applications that produce Grand Challenge scientific results, but we must also execute those applications on state of the art computational platforms. To efficiently match the current computational requirement of CT scientists with the available resources and to motivate the development of useful machines with even more capabilities, we must quantify the current codes and the testbeds on which they run. We hope that by extrapolating the requirements of the CT community we can influence the development of even more capable hardware.

This grant supported this effort to characterize the problem domain of the Earth Science Technology Office's Computational Technologies Project, to engage the Beowulf Cluster Computing Community as well as the High Performance Computing Communities so that we can predict the applicability of said technologies to the research community represented by the CT project and formulate long term strategies to provide to the computational resources necessary to satisfy the computational requirements of the the CT science community. The ultimate goal of the evaluation effort is to use the information gathered over the course of the Round-3 investigations to quantify the trends in scientific expectations, the algorithmic requirements and capabilities of high-performance computers to satisfy this anticipated need.

The Computational Technologies Project represents a slight change in direction from the preceeding rounds of this project—rounds one and two of the HPCC/ESS project. In Round-2 the evaluation effort focused on in-depth studies of the behavior of individual codes primarily on the T3E. During Round-2 a number of developments occurred that are changing the way we think about high-performance computing. The highly capable codes developed in Round-2 have exposed the fact that in addition to being high performance, the codes must also be maintainable, have interoperable components, and be interoperable at least within a science community and portable across multiple architectures.

The move to a framework approach to developing models can be seen as a natural result of "environmental" conditions. One characteristic of the CT community that

leads naturally to a frameworks approach is that many of the codes have a long history and a large community that actively participate in their development. In some sense the codes that implement the current model are the embodiment of the development of the model itself. There can be pieces of the model that are quite old and well established. They are written in the style and language of the time. Other pieces of code are the focus of current development. They are constantly being changed by a number of researchers. Managing this disparate code base is beyond conventional software project techniques.

Cluster computing has reached the performance level where it has Grand Challenge science implications. Open Source software now plays a major role in most scientific computing environments and some believe it holds key to closing “the software gap” in future high performance systems.

### **Activities to Support the ETSO CT Project**

This was a two year grant, spanning the time period from 2003/05/01 to 2005/04/30. In the first year we developed a prototype system `mprof` to do source code level instrumentation of C programs that use MPI and UPC programs. The tool used hardware counters through the `papi` interface software and the analysis tools `Pablo` and `xmpi`. During the development we used standard demo codes used in the HPCC summer school and the sequential version of the `kriging` code provided by Dr. Pedelty. We met with the PI teams in an attempt to establish the goals of instrumenting their parallel codes as they became available.

In support of the ETSO CT project, I participated as a member of the in-house team of lectures in the HPCC Summerschool 2003.

There was no funding available for the second year of the grant, so this project effectively terminated in spring of 2004. There were, therefore, no reportable research activities in the second year of this grant.

Residual funds from the first year were used to support my participation as a lecturer in the final HPCC Summerschool in July of 2004.